DOCUMENT RESUME

ED 182 574 CE 024 259

AUTHOR Espy, John: Selleck, Ben
TITLE Nuclear Technology. Course 26: Metrology. Module 27-7, Statistical Techniques in Metrology.
INSTITUTION General Physics Corp., Columbia, Md.
SPONS AGENCY Office of Education (DHEW), Washington, D.C.
PUB DATE 79
CONTRACT 300-77-0340
NOTE 66p.: Small, light type in figures will not re For related documents see CE 024 245-298

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS Course Content: Energy: Learning Activities: Learning Modules: Measurement Instruments: Measurement Techniques: *Nuclear Physics: Postsecondary Education: *Quality Control: Statistical Analysis: *Statistics: Teaching Guides: *Technical Education
IDENTIFIERS Nuclear Energy: Nuclear Energy Occupations: *Nuclear Power Plant Technicians: *Nuclear Technology

ABSTRACT

This seventh in a series of eight modules for a course titled Metrology focuses on descriptive and inferential statistical techniques in metrology. The module follows a typical format that includes the following sections: (1) introduction, (2) module prerequisites, (3) objectives, (4) notes to instructor/student, (5) subject matter, (6) materials needed, (7) procedures, (8) data table, (9) reference materials, (10) test. (The course is one of fifteen technical courses for the second year of a two-year postsecondary curriculum in Nuclear Quality Assurance/Quality Control Technology.) (LRA)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED182574

NUCLEAR TECHNOLOGY

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

COURSE 27 METROLOGY

MODULE 27-7 Statistical Techniques in Metrology

CE 024 859

GENERAL PHYSICS CORPORATION
Columbia , Maryland

COURSE MODULES

COURSE 27 METROLOGY

Module 27-1	Fundamentals of Metrology
Module 27-2	Fixed Gages, Dividers, Calipers and Micrometers
Module 27-3	Gage Blocks, Mechanical Comparators and Electronic Comparators
Module 27-4	Angle Measurement Instruments, Optical Projectors and Surface Texture Gages
Module 27-5	Tolerancing
Module 27-6	Calibration
Module 27-7	Statistical Techniques in Metrology
Module 27-8	The Total Measurement System

The material in this module was developed by General Physics Corporation under a contract with the United States Department of Health, Education and Welfare, Office of Education, Contract Number 300770340. Contractors undertaking such projects are encouraged to express freely their professional judgment in the conduct of the project. Therefore, points of view or opinions stated do not necessarily represent official Office of Education position or policy.

This module was prepared for General Physics Corporation by John Espy, General Physics Corporation, and Ben Selleck, Ben-Sel Associates.

GENERAL PHYSICS CORPORATION

1979

INTRODUCTION

The QA/QC Technician's responsibilities include the determination of acceptability or non-acceptability of all parts received at a nuclear plant or construction site. In addition to verifying that shipping, packaging and storage requirements have been met for these parts, the QA/QC Technician must determine the likelihood that parts received are going to function in a satisfactory way when put into use in the various operating systems for which they are intended. Another such responsibility is the determination that the various systems, such as a pumping system, are likely to be functioning properly. It must be known, among other things, whether or not the "dimensions" and "tolerances" are within prescribed limits for each type of part received. It must be known if a pumping system is pumping liquids at the prescribed temperature and pressures.

Given the number of parts that can be received in a single shipment (in the thousands) or the volume of output of a pumping system (in the thousands of gallons), it is unreasonable to expect an examination of every part received. It is unreasonable to expect continuous measurement of a system's output. Therefore, the QA/QC Technician needs to know how to make decisions about the acceptability of a total shipment of parts by examining some representative sample of these parts. She/he must make a decision about a system's output from a sample of measurements of the system's output. To accomplish these tasks, it is essential to recognize and apply basic statistical concepts that relate data gathered from a sample of objects to predictions about the entire set of objectives.

It is extremely important that such samples of objects be *random samples*. A sample is said to be random when every object in the entire set of objects has the same chance of being selected for the sample as every other object in the

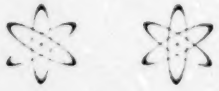
entire set. If the sample is not random, then any prediction based on the sample is likely to be in error. The sample is then said to be *biased.* In the mid 1930's, a poll was taken of a sample of voters to predict the winner of the 1936 presidential election. The results of the poll indicated that Alf Landon would win by a landslide. Of course, history records Franklin D. Roosevelt as the overwhelming choice of the public. The problem with the prediction was a biased sample. Names were chosen at random from telephone directories and state motor vehicle records. However, telephones in the 1930's were most likely to be possessed by upper middle class and upper class whites, and automobiles were owned predominantely by older, white, upper-middle class and upper class males. These people were most likely to be republican. Middle and lower class, white collar and blue collar workers who were likely to be democrats did not have the same chance to be selected for the sample as did the upper-middle class white male republicans. The sampling plans presented in this module employ procedures for selecting truly random samples so as not to be subject to errors introduced by sampling bias.

Statistical techniques are employed in metrology in three important ways:

1. descriptive statistics, such as the average size or weight of an item, are used to define the characteristics;
2. inferential statistics are used as a basis for sampling some items from a batch of manufactured goods and deciding if the entire batch should be accepted or rejected;
3. assessing the accuracy of the measuring instrument itself.

This module will concentrate upon the first two uses, but it should be noted that the way a person measures an object always affects the accuracy of the results. It is not unusual for two people to come up with slightly different numbers when measuring the same object. Statistical techniques covered in points 1 and 2 above will provide the QA/QC technician with a method for quantifying the measurement variations that result from the equipment used as well as from the person making the measurements.

Often the patterns of variation look very much like the *Gaussian distribution curve*, commonly called the *"normal curve"*, which is probably the most utilized tool for anyone making a statistical analysis. It has been used, with varying degrees of justification, to describe the distribution of the heights of males and of females, the sizes of beans, the weights of newborn babies, the velocities of particles in a gas, the number of chocolate chips in cookies, and numerous other properties of the physical and biological worlds. The normal curve will figure quite prominently in some of the quality control work presented in this module.



MODULE PREREQUISITES

A Basic Algebra Course is essential for this module. The student/trainee shall also have completed all previous modules of this course.

OBJECTIVES

Upon completion of this module, the student/trainee can do the following:

1. Describe the general responsibilities of the QA/QC technician with respect to quality assurance, quality control and process control at a nuclear plant or construction site.
2. Given a set of measurements, identify the appropriate *measurement cells* and their corresponding mid-points or mid-values.
3. Given a set of measurements construct a *frequency distribution* (tally sheet) of the measurements using appropriate measurement cells.
4. Given a set of numbers, compute the *mean* and the *standard deviation* using the formulas:

   $$\text{Standard deviation} = s = \sqrt{\frac{N\Sigma fx^2 - (\Sigma fx)^2}{N(N-1)}}$$

   $$\text{Mean} = \bar{x} = \frac{\Sigma x}{N}$$

5. Describe three quality acceptance *sampling* schemes and state the weaknesses and strengths of each scheme.
6. Describe the risks associated with sampling procedures.
7. Given a set of sampling parameters and MIL-STD-105D, identify the appropriate sampling size and the acceptance and rejection *criteria*.

NOTES TO INSTRUCTOR / STUDENT

Some of the formulas and procedures presented in this module were chosen to facilitate the use of pocket calculators -- in particular the formula for the *standard deviation** of grouped data. It is recommended that the student/trainee be given instructions if necessary and the opportunity to practice the use of the pocket calculator. If you (the instructor) prefer a more sophisticated treatment of the standard deviation, then any suitable statistics text may be used as a supplement.

An appropriate lab exercise that you might use to summarize and review the content of the module if it is possible at your location is as follows: Provide the student/trainee with a single box of nails which might represent a shipment received. Have each student/trainee select a sampling plan and make the appropriate measurements on the sample with a micrometer. The student/trainee will then perform all appropriate calculations, and with the use of MIL-STD-105D, accept or reject the lot. The measurement data obtained by each student/trainee could then be polled and the precision of the measurements could be determined.

*NOTE: All italicized words in this module are defined in in the glossary of terms.

SUBJECT MATTER

GLOSSARY

- AQL - Acceptable Quality Level (associated with Producer's Risk).
- Acceptance inspection - a process for estimating the qualities of a group of items called a "lot" and then deciding to accept or reject the lot.
- Acceptance number - cut-off level set by sampling plan that specifies the maximum number of defectives that can be found and still have the lot accepted.
- Attribute data - refer to measurements collected with regard to whether they are within or beyond same specified requirement.
- Biased sample - a group of measurements that are <u>not</u> randomly drawn from the population.
- Consumer's risk - probability that a bad lot will yield an acceptable sample.
- Cell limits (boundaries) - are the extreme upper and lower measurements allowable in a measurement cell.
- Central tendency - most commonly refers to the arithmetic mean that is found by summing the sample measurements and dividing by the number of measurements taken.
- Criteria - a standard agreed upon for statistical evaluation purposes.
- Deviation - dispersion of measurements about a central tendency.
- Frequency - is the number of times (f) that a particular measurement (x) occurs in a measurement cell.
- Frequency distribution - a graph of the number of measurements observed within each of the measurement cells.

- Gaussian (normal) curve - a frequency distribution (e.g., of sample measurements) that looks like a "bell" shape.
- Interval - units of measurements on a scale.
- Interval of uncertainty - the unit on a scale that spans the mid-point of a measurement cell and within which, measurements are considered equal to each other.
- Lot - collection of discrete units.
- LTPD - Lot Tolerance Percent Defective (associated with Consumer's risk) - level of defectiveness that is unsatisfactory and thus should be rejected by the sampling plan.
- Mean - is the average of a set of measurements found by summing the measurements and dividing by the number of measurements.
- Measurement cell - represents an interval of uncertainty concerning true measurements.
- Mid-point (mid-value) - of a cell is the value obtained by adding the cell limits and dividing by 2.
- MIL-STD-105D - Military Standard (see resource list at end of module).
- Normal curve - see Gaussian distribution curve.
- Operating - characteristic curve (OC curve) - graphically represents the tradeoff between the chance of acceptance of a lot and the quality of the lot. (See Figure 27-7-8).
- P_a - probability of lot acceptance.
- P_r - probability of lot rejection.
- Probability - likelihood of an event taking place (usually expressed as a ratio).
- Process control - involves the taking of successive samples of output of a system and applying statistical

techniques to determine if the system output is under control.

- Producer's risk - probability that a bad lot will yield an acceptable sample.
- Random samples - a group of measurements made on items (events) drawn from a population (lot) in a non-biased or equal-probability fashion.
- Range - the measure of variability that is obtained by taking the largest measurement minus the smallest measurements, the difference being the range.
- Sampling plan - a method for taking measurements on a sample drawn from a lot (or process) and the statistical treatment of the measurements in order to accept or reject the lot (or process control).
- Standard deviation - a measure of dispersion of a set of numbers about the mean (see formula in Central Tendency section).
- Summation - symbolized by a greek sigma (Σ), denotes a list of measurements (x's) to be added together.
- Variable data - measurement data that represent specific or exact values.

DATA COLLECTION

Collecting data can be considered detective work -- "numerical detective work" -- or "graphical detective work". In the case of the QA/QC person, clues are being sought as to the character of shipments received. She/he is looking for information concerning the reliability or usability of parts. The QA/QC person, however, is more than a detective. She/he is also judge and jury. She/he collects data, organizes and evaluates data, and determines the fate of the shipment (accept or reject).

The first task after measurements are taken and collected is to write down the numbers in such a way that a general idea of "what they are like" is presented. Some examples of such groups of measurements are (a) the average speed of each car in a particular race, (b) the batting averages of baseball players on a certain team, (c) the number of ads for QA/QC Technicians in all of the newspapers in each of ten major cities in the United States, or (d) the heights of fifteen 10th graders in a particular homeroom. There are many ways to write down these sets or groups of numbers.

The heights of the 10th graders, for example, may be written as they were taken: 5' 9.5", 5' 6", 5' 11.5", 6' 3", 5' 11", 5' 9.5", 5' 11", 5' 6", 5' 6.5", 5' 11", 5' 11", 6' 2", 5' 9", 5' 8.5", 5' 11.5".

This arrangement does give some idea of what the heights of the 10th graders are like but there are ways to write these numbers to get a clearer picture. For example, they may be written in <u>order</u> from shortest to tallest: 5' 6", 5' 6", 5' 6.5", 5' 8.5", 5' 9", 5' 9.5", 5' 9.5", 5' 11", 5' 11", 5' 11", 5' 11", 5' 11.5", 5' 11.5", 6' 2", 6' 3".

With this arrangement, it can be seen at a glance that the heights of the 10th graders range from 5' 6" to 6' 3". It is a better picture of "what they are like" than is the random picture first presented above; however, another picture may be presented by using a chart that shows the <u>frequency</u> of each of the heights as shown in Figure 27-7-1.

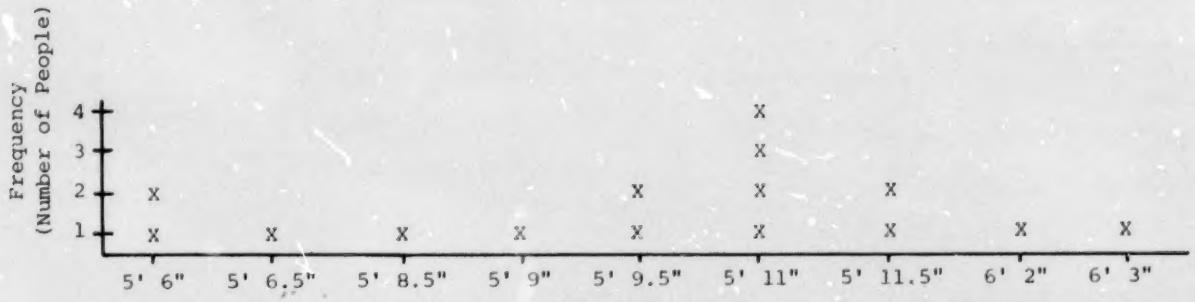


Figure 27-7-1 Frequency Distribution According to Height

With this arrangement it can easily be seen that the *range* of heights is from 5' 6" to 6' 3", and in addition it can easily be seen that four 10th graders are 5' 11", two are 5' 11.5", two are 5'6". What is not readily obvious from this arrangement, however, is that no 10th grader is 5' 7" or 5' 8" or 5' 10" etc. If a chart were used that showed all possible heights in half inches from 5'6" to 6'3", this information would be readily available by simple inspection:

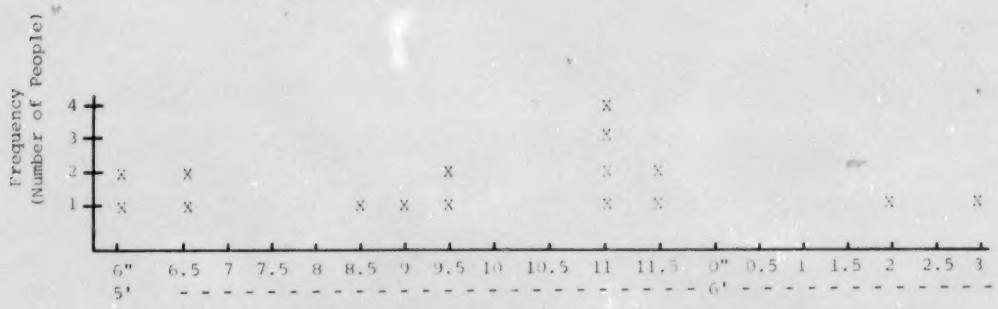


Figure 27-7-2 Frequency Distribution According to Height Shown in Half-inch Increments.

Measurement Cells

It should be noted that no measurement is absolute but only an approximation. In the examples above, it would be accurate to have described the measurements as approximately 5' 11" or approximately 6' 2" etc. How approximate depends on the measuring device that is used. (See Module 24-4). For any measurement that is made, other than a frequency count, there is always an *interval of uncertainty*. This interval depends upon the scale used on the readout device. For example, if the measuring instrument is centimeters (see Figure 27-7-3), the measures of each of the segments $\overline{AB}$, $\overline{CD}$ $\overline{EF}$, and $\overline{GH}$ would be reported as 3 centimeters to the nearest centimeter. The line segment, $\overline{MN}$ and $\overline{OP}$ would be reported as 8 centimeters to the nearest centimeter. For $\overline{AB}$, $\overline{CD}$, $\overline{EF}$, or $\overline{GH}$ all we can be sure of is that the true measure is between 2.5 and 3.5. This is an interval of uncertainty. For $\overline{MN}$ and $\overline{OP}$ this interval of uncertainty is from 7.5 to 8.5. It is possible to make use of these intervals of uncertainty and construct something called measurement cells.

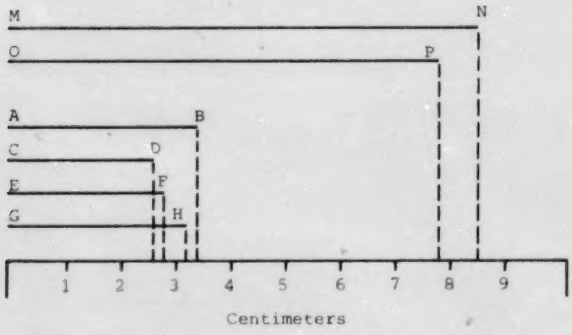


Figure 27-7-3 Measurement Cells Denoting an Interval of Uncertainty

Taking the average of the measures that group together and halving the interval of uncertainty yields a statistical tool called a measurement cell.

Measurement cell = average value $\pm$ ½ (interval of uncertainty)

In Figure 27-7-3 we would then have two measurement cells at 3 $\pm$ 0.5 and 8 $\pm$ 0.5.

In Figure 27-7-4, the interval of uncertainty is stated as $\pm$0.5 about each average measurement in centimeters and inches.

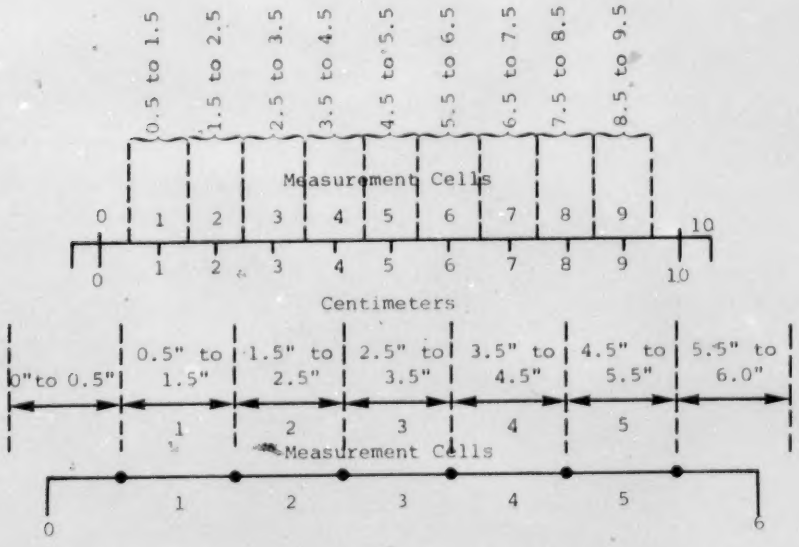


Figure 27-7-4 Measurement Cell Example

Even though no measurements are absolute, they can be treated as absolute in terms of the measurement cell to which they belong. The value assigned is the value of the midpoint of the interval to which its measurement belongs. In Figure 27-7-3, the value assigned to $\overline{AB}$, $\overline{CD}$, $\overline{EF}$, and $\overline{GH}$ is 3. 3 is the midpoint of the cell to which all of these measurements belong.

In general, the *midpoint* or *midvalue* of a cell is the value assigned to all measurments that fall within the cell limits or boundaries. *Cell limits* or *boundaries* are the extremes within which measurements must lie if they are to be assigned the value of the midpoint. In Figure 27-3-3, these cell limits or boundaries for the measures assigned $\overline{MN}$ and $\overline{PO}$ are 7.5 and 8.5.

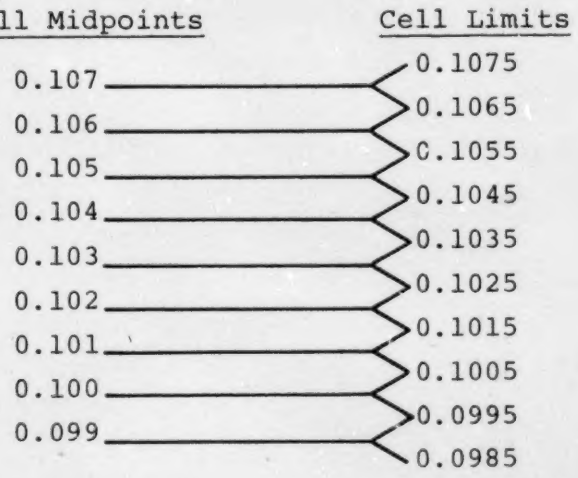
For tolerance measurements, if parts are being measured to the nearest 0.001 inch and they run approximately 0.1000 inch, the cell for this value would be 0.1000 as the midpoint and 0.0995 as the lower cell limit and 0.1005 as the upper cell limit, or 0.1000 ± 0.0005. This could be shown as follows:

Cell Midpoint	Cell Limits
0.1000	0.1005
	0.0995

Note that this means that any reading greater than 0.0995 and less than 0.1005 would be considered 0.1000. Assuming a cell interval of 0.001 and an average size of 0.103, the following would be a way to tabulate the data collection:

Table 27-7-1

Cell Midpoints and Cell Limits Table



Cells could be added or deleted from each end as needed.

Once the measurement cells have been established, data can be collected by tallying (or counting up) the measurements that fall in their appropriate cells, using the midpoint as the value.

NOTE: Cell intervals need not be in units of one. However, once established, a cell interval must be constant for a given distribution (tally).

CENTRAL TENDENCY

The writing down of sets of numbers to get a good idea of what "they are like" is very useful to the numerical detective, but there are characteristics of these sets that are not apparent for any display of the numbers themselves. Of interest to the statistician, among other things, concerning groups of measurements are the *central tendencies* of

these numbers. Central tendencies are determined by arithmetic computation quite often. One such measure of a central tendency is the arithmetic *mean* (denoted by $\bar{x}$), which is computed by summing all of the measurements in the group and dividing that sum by the number of measurements in the group. For example, consider the following group of 7 numbers: 15, 19, 23, 12, 27, 35, 16.

The formula for the mean is:

$$\bar{x} = \frac{\sum_{i=1}^{k} X_i}{N}$$

In the formula above, $\bar{x}$ = the mean or average of the measurements x, Σ = summation of all x values, from subscript i = 1 to i = k (all x values here), N = number of measurements summed.

The calculation proceeds as follows:

$$\bar{x} = \text{The mean} = \frac{15 + 19 + 23 + 12 + 27 + 35 + 16}{7}$$

$$\bar{x} = \frac{147}{7}$$

$$\bar{x} = 21$$

The measures of central tendency, such as the mean, do provide more information about what the measures look like, than just the displays of numbers; however the mean does not tell very much about how the measures disperse themselves around the mean.

In the following example, Figure 27-5-5, 18 people's scores on a test have a mean of 50.

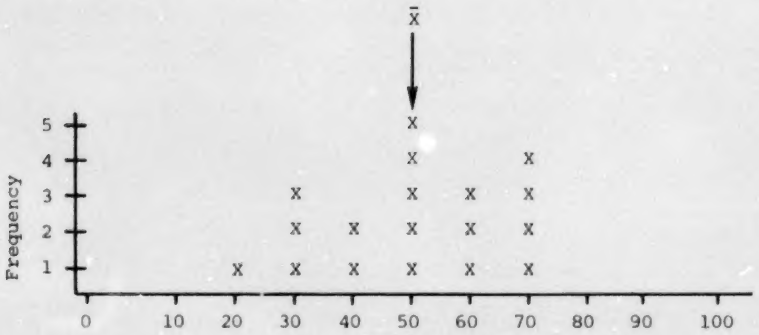


Figure 27-7-5 Frequency vs. Score Graph Displaying Central Tendency

The scores range from 20 to 70. They cluster around 50 (the mean).

Another group of 18 scores for the same test might also have a mean of 50 with a wider spread (dispersion) of scores. See Figure 27-7-6.

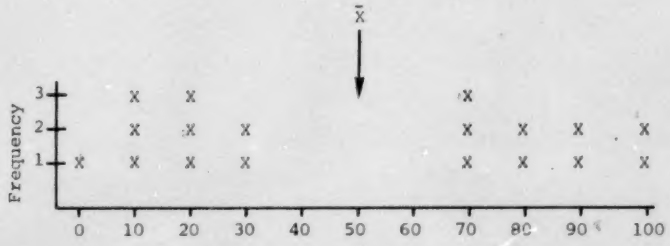


Figure 27-7-6 Frequency vs. Score Graph Showing Wider Dispersion of Data

These 18 scores range from 0 to 100. They do not cluster near their mean of 50, but are more widely spread out from the mean.

In the example shown in Figure 27-7-5, the measures *deviate* very little from the mean while in the example in Figure 27-7-6, the measures deviate quite a bit from the mean. These deviations give the numerical detective (the QA/QC person) more information and therefore, an even better picture of what the measures are like than does a tally sheet.

They give a picture of how the individual measures are dispersed about the mean of the measures.

The Standard Deviation

The most widely used measure of the dispersion, or scatter, of scores about their mean is the *standard deviation*. As has been seen, the mean value of any set of measurements (for example, the lengths of a bolt) can always be computed. However, this gives no information about the likely uncertainty in subsequent measurements. Knowledge of the standard deviation for a set of measurements allows us to make a statement such as "it is 50% probable that a future measurement of the bolt length will lie between some upper limit (say 6.2 inches) and some lower limit (say 6.0 inches). These limits are established by the value of the standard deviation, with a larger value signifying a large uncertainty and vice versa.

The formula for the *standard deviation* is:

$$s = \sqrt{\frac{N\Sigma fx^2 - (\Sigma fx)^2}{N(N-1)}}$$

Where

- s = standard deviation
- $\sqrt{\ }$ = square root
- N = number of all measurements taken
- Σ = the sum of
- f = frequency of scores in measurement cell
- x = midvalue of measurement cell

The formula will provide a measure of dispersion that is not available from calculating the mean alone.

An illustration of the use of a data table to determine the standard deviation of a set of measurements is presented below:

Measurement Cell	Mid-Point (x)	Frequency (f)	$f \cdot x$	x^2	$f \cdot x^2$
14-16	15	1	15	225	225
12-14	13	2	26	169	338
10-12	11	4	44	121	484
8-10	9	3	27	81	243
6-8	7	1	7	49	49
Sums		$N = 11$	$\Sigma fx = 119$		$\Sigma fx^2 = 1339$

The standard deviation, s, is:

$$s = \sqrt{\frac{N\Sigma fx^2 - (\Sigma fx)^2}{N(N-1)}}$$

$$s = \sqrt{\frac{11(1339) - (119)^2}{11(10)}}$$

$$s = \sqrt{\frac{14729 - 14161}{110}}$$

$$s = \sqrt{\frac{568}{110}}$$

$$s = \sqrt{5.164}$$

$$s = 2.272$$

The mean value, $\bar{x}$, is:

$$\bar{x} = \frac{\Sigma f \cdot x}{N}$$

$$\bar{x} = \frac{15 + 26 + 44 + 27 + 7}{11}$$

$$\bar{x} = 10.818$$

THE NORMAL CURVE

The *standard deviation* is most frequently used in connection with something called the *normal curve*. It was stated in the introduction to this module that the patterns of variations in every aspect of measurement often look like a normal curve. This is particulary true when the measurements represent a large number of objects or outcomes. In these instances, most of the measurements tend to cluster near the mean with fewer measurements found at the extremes. To illustrate this, consider the distribution of scores represented in Figure 27-7-5 and figure 27-7-6. If the columns of x's are united by a curve as in Figures 27-7-7a and b, the curves that are formed are difficult to describe.

In the first case, the measurements are all close to the mean with none at either extreme. In the second case, all measurements are at one extreme or the other with none near the mean. When most of the measurements are near the mean with a few at both extremes the curve takes on the appearance of a bell. Please refer to Figure 27-7-8a and b. When this happens, the curve that is formed is called a normal curve and the measurements are said to be normally distributed or dispersed about the mean.

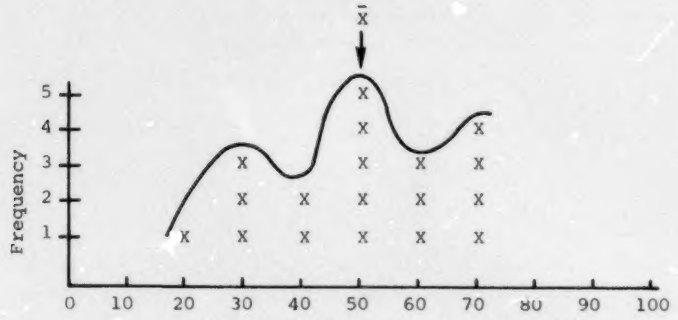


Figure 27-7-7a Frequency vs. Score Graph Showing Central Tendency

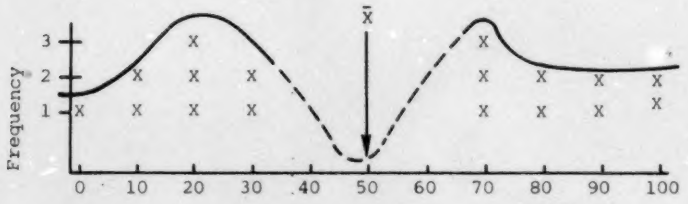


Figure 27-7-7b Frequency vs. Score Graph Showing Wider Dispersion of Data

Area Under the Curve

In normal distribution of measurements, the standard deviation of the measurements is used to determine the area under the normal curve. The standard deviation is commonly used as the measurement unit. The total area under the curve is said to be 100%.

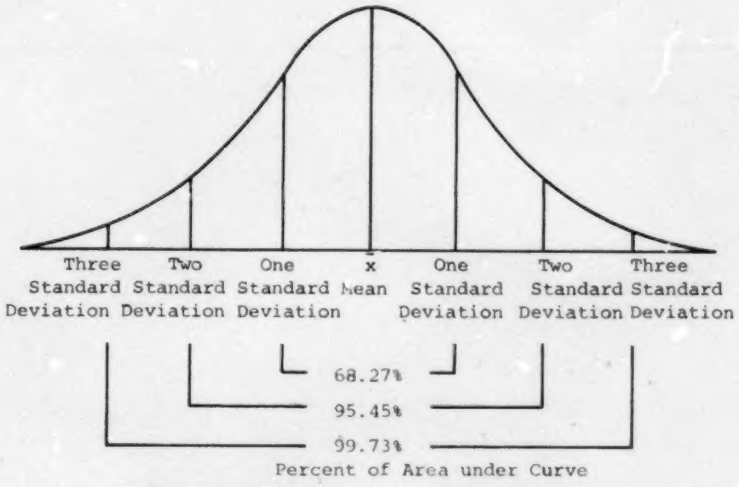


Figure 27-7-8a Theoretical Normal Distribution

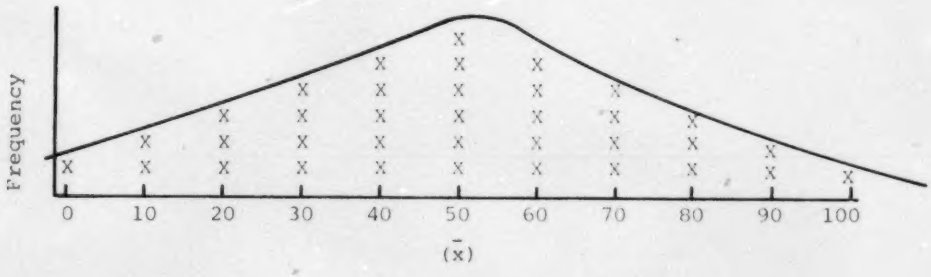


Figure 27-7-8b Observed Normal Distribution

In a normal distribution, 34.135% of all measurements are within an interval of one standard deviation above the mean; another 34.135% are within an interval of one standard deviation below the mean. This means that 68.27% (34.135% + 34.135%) of measurements are found in a central interval from one standard deviations below the mean to two standard deviation above the mean. 95.45% are in an interval from two standard deviation below the mean to two standard deviations above, and 99.73% will be in an interval from three standard deviations below to three standard deviations above. See Figure 27-7-8a.

Uses of the Normal Curve

The QA/QC Technician can use the properties of the normal curve in making decisions in both acceptance inspection and process control operations. For example, if a supplier of screws claims that the mean of the measurements of the screws' lengths are 2.5 cm. with a standard deviation of $\pm$ 0.05 cm., the QA/QC Technician can use this information to determine the expected range of measurements of the lengths of the screws in the shipment. Thinking in terms of the normal curve (See Figure 27-7-9) the technician knows:

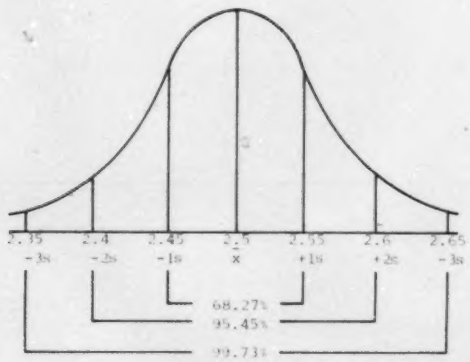


Figure 27-7-9. Expected Normal Distribution for Screw Measurements.

1. 68.27% of the screws should measure between 2.45 cm. and 2.55 cm. - the limits of the central interval (±1 standard deviation).
2. 95.45% of the screws should measure between 2.4 cm. and 2.6 cm. (±2 standard deviation).
3. 99.73% of the screws should measure between 2.35 cm. and 2.65 cm. (±3 standard deviations).
4. Only 3 screws in every 1000 should be either longer than 2.65 or shorter than 2.35 cm.

Now when the QC/QA Technician measures the lengths of a representative sample of screws from the shipment, he/she knows what measurements to expect if the mean is really 2.5 cm. (that is if all screws could be measured) with a standard deviation of 0.05 cm. If the measurements of the sample are "too much" different from what is expected, the technician has strong evidence to suspect that the supplier's claims about the mean and/or standard deviation of this shipment of screws is not valid. This kind of information is valuable in deciding whether to accept or reject the shipment.

The Exercise and Procedures sections of this module give problems with solutions which reinforce the concept of THE NORMAL CURVE. These should be worked to ensure proper understanding of this material.

TYPES OF DATA

There are basically two types of measurement data - variable and attribute - that the QA/QC people must deal with. *Variable data* are measurement data that represent specific values. The examples given previously in the

section on DATA COLLECTION cited that a) the average speeds of each car in a specific race, b) the batting averages of baseball players on a certain team, c) the number of ads for QA/QC positions in newspapers and d) the height of fifteen 10th graders, are all examples of variable data. *Attribute data* are data that indicate whether measurements are within or beyond some specified requirement. For example, a sample of bearings are being examined by a type of tolerance gauge where those that do not pass through are rejected. Specific measurements of each bearings are not made, only a measurement of whether or not each bearing is within or beyond some specific requirement. It should be noted that whether data is variable or attribute really depends on how the data is used. Specific measurements may be taken but the data could still be attribute. For example, the Military Academy at West Point will not accept applicants who are over 6'6" tall. Specific measurements of heights are made but they are classified as under 6'6" or over 6'6". They are either within or beyond a specified requirement. They are attribute data.

In many instances at receipt inspections, the QA/QC Technician may find it impractical or not necessary to secure specific measurement data. It may be desirable to consolidate a number of quality characteristics such as those of an assembly and simply define the entire unit as conforming or non-conforming. These kinds of data are normally expressed as either percent defective or rejected of the pieces inspected and/or tested. However, on occasion, the QA/QC will deal with the variable data, particularly with respect to something called *process control*. This will be discussed more fully later in this module.

STATISTICS AND REPETITIVE PROCESSES

All of the previous ideas come together in systematic statistical quality control, which is the statistical surveillance of repetitive processes. Any repetitive process may call for quality control. For example, turning out rivets, building coolant pumps, generating electricity, receiving purchase orders, or recycling water involve a large volume of activity composed of distinct units repeated many times. The purposes of statistical quality control are (1) ACCEPTANCE INSPECTION, which tries to evaluate past performance, and (2) PROCESS CONTROL, which attempts to estimate present and future performance. In both cases decisions are made on the basis of samples taken. On the basis of an analysis of the sample, in process control it might be decided whether a process is still under control while for acceptance inspection an estimate of the qualities of a group of items, called a "lot", is made and the lot is accepted or rejected. Most of the QA/QC person's time in statistical surveillance of repetitive processes will be spent in acceptance inspection at the times of receipt inspections. However, he/she will have occasion to become involved in process control and therefore both concepts are presented in this module.

ACCEPTANCE INSPECTION

The QA/QC person might receive a shipment of detonating fuses and have the responsibility for designating the entire lot as "acceptance" or "unacceptance". Assuming that the installation can tolerate no more than 2% of those fuses as duds, how can it be fairly ascertained that the entire shipment meets those specifications? Obviously, all of

the fuses cannot be tested no matter how large or small the shipment size, as testing destroys them. Therefore, only a sample is tested. If the sample contains a specified number of defective fuses or more, the entire lot is rejected -- otherwise it is accepted.

Even if testing did not destroy the items being inspected, it may not be desirable to test 100% of them. Inspector fatique, among other things would interfere with the accuracy of the results. According to Juran (Quality Control Handbook, p. 12-51) 100% inspection at best is only correct 80% of the time. Sampling procedures reduce the opportunity for inspection error in individual measurements.

Such a sampling plan has two risks: first, there is the risk that a good lot will yield a bad sample; and second, the risk that a bad lot will yield a good sample. Now the first risk is not really the problem of the buyer, as the good lot which tested bad would be taken back by the producer. The risk is generally termed the *producer's risk*, or the seller's risk.

The second risk is the problem of the buyer. A small sample could be tested, and just by the luck of the draw be a good sample which would lead to acceptance of the entire lot. This risk is termed the buyer's risk, or the *consumer's risk*. This risk should be kept as low as possible. One way to do so would be to check if the shipment included a certified control chart covering this lot. If the control chart were available, the QA/QC inspection could be limited possibly to only an audit of the supplier's control charts.

The second way to reduce this risk would be to use a sampling plan which gives an advantage to the buyer. Certain plans are more likely to assist the inspector in rejecting most of the bad lots that may be submitted, but

they may increase the number of good lots mistakenly considered bad. Of course price might increase considerably if too many lots are rejected so both the buyer and the seller have an interest in both types of risks.

A variety of sampling plans exist - separate, well-thought-out procedures with clear rules for choosing a sample of a certain size in a certain way, and rejecting the lot if the number of defective items exceeds a specified number. With any given one, the acceptance or rejection of a lot is a matter of chance since it depends on a random sample. However, the probability of acceptance depends on the true quality of the lot -- which leads to the common saying that "quality cannot be inspected into the product, it must be built in."

The tradeoff between the chance of acceptance of a lot and the quality of the lot can be represented by an *operating-characteristic curve*, usually called an *OC curve*. Most OC curves resemble the Figure 27-7-10. The QA/QC Technician would accept a smaller and smaller percentage of the lots as poorer and poorer quality is considered. Obviously if the lot has zero percent defects, it is certain to be accepted; hence the OC curve always starts from 100. On the other end, if the lot has 100 percent defectives, it would certainly be rejected and the OC curve is at zero for that level of quality. The curve represents what a particular sampling plan will accomplish, and therefore it is of vital concern to the technican who must carry out the sampling and decide on the acceptance of a particular lot.

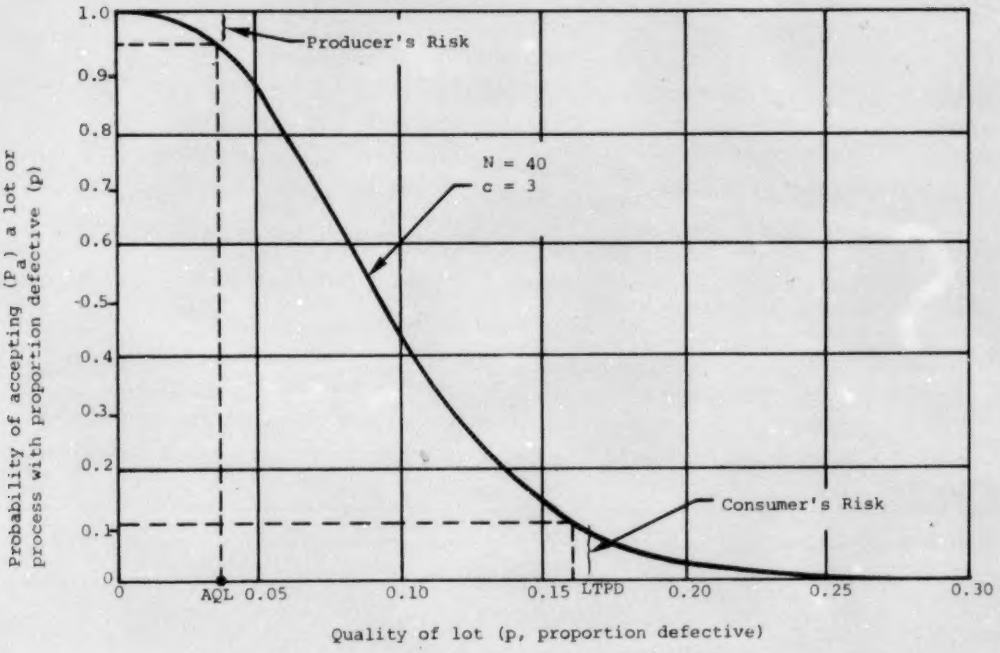


Figure 27-7-10 Operating-characteristic Curve for a Typical Sampling Plan Having a Sample Size = 40 and Acceptance Number = 3.

The points on the horizontal scale are possible lot quantities (i.e., proportions (P) defective), and the height of the curve shows the probability that a particular lot of this quality will be accepted (P_a) under a specified sampling plan. In the figure above, the sampling plan calls for accepting the lot if three or fewer defectives are found in a sample size of 40.

Thus if a lot that is 5 percent defective (0.05 on horizontal line above) is submitted to the plan whereby the sample size, N, is 40 and the acceptance number (c) is 3, there is slightly less than a 0.90 probability that the lot would be accepted. Note that one defective beyond the acceptance number calls for rejection of the lot.

Often the table is read in reverse fashion from the last interpretation. The *acceptable quality level* (AQL) is the level of quality at which the probability of acceptance of a lot is fairly large, typically around 0.95. For example in the illustration given, the AQL would be about 3 percent (0.03) if a 0.95 probability of acceptance were used.

Those same figures could be looked at another way. From the producer's point of view there is the risk that the material, whose quality is at the AQL, would be rejected by the acceptance plan. In the illustration, the producer stands a risk of 0.05 (1.00 - 0.95 probability of acceptance) that material that is 3 percent defective would be rejected.

At the other end of the curve is shown the consumer's risk, which is the risk that the user of the material takes that a lot of material actually having an undesirable quality level will be accepted by the plan. For example, in the figures above, the consumer is taking slightly over a 0.10 chance that material which is actually 15 percent (0.15 as as a proportion) defective will be accepted.

Sometimes the consumer risk is specified as a relatively low probability, generally 0.05 or 0.10. In the military standard entitled, "Sampling Procedures and Tables for Inspection by Attributes (MIL-STD-105D)," the term *Limiting Quality (LQ)* is used to refer to the lot quality associated with a particular consumer risk chosen. A second term related to LQ is the *Lot Tolerance Percent Defective (LTPD)*. The LTPD refers to the lot quality associated with the specific

probability of acceptance level of 0.10 (consumer risk of 0.10). In the example given in Figure 27-7-10 a LTPD of approximately 0.16 is observed (for a probability of acceptance (consumer risk) of 0.10.

Each sampling plan has a different OC curve, and the QA/QC Technician often uses different plans for different items being inspected. An ideal sampling plan would accept all good lots and reject all bad lots without error. But such a plan would require inspection of all, or practically all, of a lot. Assume for instance that two percent defectives or fewer is acceptable quality and over two percent unacceptable for a lot of 400,000 widgets. If one had taken a sample of 399,999 and found 8,000 defectives one would still not be absolutely certain that the lot was not over two percent defective -- the final item might be defective and 8,001 is more than two percent of 400,000. Therefore, unless some uncertainty in the immediate vicinity of the cutoff point is tolerated, complete inspection is necessary.

And even with 100 percent inspection there would not be certainty in practice of accepting all good lots and rejecting all bad lots. Even top-notch technicians and their instruments are fallible. A tendency exists to be less careful if a sample is large than if it is small. Worse yet, after a technician has looked at a few thousand items, they all begin to look alike. But the steepness of the OC curve depends on the sample size, and the steepness is the measure of the discrimination or sharpness of the test of acceptability. In many instances, the requirement for greater sharpness in differentiating between three and six percent defectives for given risks, for example, which might require a very large sample to be inspected, is balanced against the psychological advantage of a smaller sample with more careful attention by the technician. Of

course 100 percent inspection makes little sense with destructive type testing such as explosive detonation, tensile, stress, rupture, or bursting type, as well as testing in which contamination of the material occurs with handling.

All sampling is based on randomness. For a truly random sample, all pieces within the lot must have an equal opportunity to be selected in the sample. Special care must be taken to insure randomness. Generally devices such as random number tables, dice, or playing cards are used to assure true randomness.

Four illustrative sampling inspection plans (see MIL-STD-105D) will show some of the variety encountered in practice. They are all based on an acceptable quality of 0.5 percent or fewer defective items in a lot.

- Single Sampling - Draw a random sample of 75 items and classify each one as defective or nondefective. With one or fewer defective, accept the lot; with two or more defective reject the lot.
- Double Sampling - Draw a random sample of 50, and classify them. Accept the lot if none is defective, reject the lot if three or more are defective. Otherwise (either one or two defectives) draw a second random sample, this time of 100 items. Now accept the lot if both samples together (150 items) contain two or fewer defectives, and reject it if the two contain three or more defectives.
- Multiple Sampling - In this type, double sampling is repeated a great many times until a final decision can be reached. As each additional random sample is taken, the combined number of defectives up to that point is used to either make a determination of accept-reject, or to continue.

- Sequential Sampling - This plan carries multiple sampling to the limit by using successive samples of one and comparing the accumulated number of defectives with acceptance and rejection numbers determined by formulas.

The OC curves of the sampling plans above are practically alike. A choice among them can be based on practical considerations, mainly costs of administration and actual inspection. The single sampling plan, the most commonly used, is easiest to administer with one sample, one recording, and one decision. The double plan has the advantage that on the average fewer pieces will need to be checked before a decision is reached. However, it is somewhat more difficult. The multiple plan requires even fewer pieces to be checked, on the average, but is still more complicated to administer. And as expected, sequential sampling is even more complicated to administer. And as expected, sequential sampling is even more economical on number of items tested -- sequential sampling commonly requires a third to a half fewer observations, on the average, than single sampling.

Usually the objective is to minimize the sample size in order to save the time required to reach a decision on the measurement in question. Generally a single sampling is the best choice. This is followed by double, multiple and the sequential sampling plans previously described.

There is no single perfect sampling plan, and there are many factors which should be evaluated to examine the variety of tradeoffs with exist. Some of the factors that have already been mentioned and additional factors that should be considered are:

1. What should be the acceptable quality level? Obviously zero defectives are desired and should always be the

goal of any quality control program. What level of quality can be sustained? While this consideration and its decision is not generally within the province of the QA/QC Technician, she/he should understand the nature of the factors going into the decision making process.

2. What are the resultant costs if different AQL's, producer's and consumers' risks are sought. How many more tests need be performed and what will they tell us? The OC curve shows the risks involved for given incoming quality levels under one sampling plan. Curves can be developed for any sampling scheme using tables commonly found in statistical quality control texts. From these curves the appropriate sample sizes and risks can be determined. Estimates of costs can then be found.
3. What is the level of capability required for acceptance tests? Does such capability exist either within the organization, or must an outside vendor be used?
4. What will be the impact of the decisions on having the material available for use at the time and place needed? To use rejected material, it may have to be altered, reworked, sorted or changed in design, or other parts and processes may have to be modified.

Use of MIL-STD-105D

The final concept to be considered in this section on Acceptance Inspection is the use of MIL-STD-105D in selecting and evaluating sampling plans. Many of the plans used in industry, such as in receiving inspection activities, are extracted from this publication and copied into company manuals, forms, and instructions. The publication contains, in addition to the sampling tables, many definitions and procedural elements, and includes operating-characteristic

curves and tables detailing the mathematical probabilities of outcomes of the various plans. This publication should be available to the student/trainee and instructor.

The publication "Military Standard - Sampling Procedures and Tables for Inspection by Attributes" (MIL-STD-105D) details the methods by which a sampling plan is selected and provides many options. Given information about the proposed sampling plan, the student/trainee should be able to identify and list various acceptance and rejection numbers, limiting quality, etc. To illustrate this process, consider the following with reference to MIL-STD-105D:

1. Sampling Inspection Plans to choose from - Single, Double, Multiple.
2. Acceptance Standards to chose from - Normal, Reduced, Tightened.
3. Lot Size - which is the number of units of product in a lot.
4. Inspection Level - I, II, III, 5-1, 5-2, 5-3, 5-4 which determines the relationship between the lot or batch size and the sample size.
5. Acceptable Quality Level (AQL) - which is the maximum number of defects per hundred units considered satisfactory.

In this example, select items for sampling as follows:

EXAMPLE:

GIVEN:

1. Sampling Inspection Plan - Single
2. Acceptance Standard = Normal
3. Lot Size = 1500 pieces
4. Inspection Level - II
5. Acceptable Quality Level (AQL) - 4.0%

DETERMINE:

The following by reference to MIL-STD-105D:

A. Code Letter Applicable - This is determined by reference to TABLE I - Sample Size Code Letters - for a Lot Size of 1201 to 3200 and an Inspection Level II, the Code Letter will be "K".

B. Sample Size and Acceptance and Rejection Numbers - This is determined by reference to TABLE II-A - Single Sample Plans for Normal Inspection (Master Table). In the Sample Size Code Letter column see "K" and read across to the 2nd column called "Sample Size", the sample size of 125. Then read under the Acceptable Quality Level (normal inspection) column labeled "4.0", the acceptance number (AC) = 10, and the rejection number (RE) = 11.

C. Limiting Quality for which P_a (probability of acceptance) is 10 percent - This is determined by reference to Table VI-A- Limiting Quality (in percent defective) for which P_a = 10 percent. Under the "code letter" column, see code K, and read across to the AQL column "4.0" to obtain the LQ of 12. This table indicates the number of rejects (12 in this case) which are accepted by the sampling plan based on a 10% probability of acceptance level (consumer risk).

NOTE: The row of numbers on the same line as the code letter K in TABLE VI-A is obtained from a set of OC curves corresponding to the sample size code K in TABLE X-K (Chart K and TABLE X-K-1) on page 48. In Chart K (the right hand chart), we see a curve labeled with the AQL of 4.0. Reading across from the P_a - 10% in the chart, an LQ of 12 is obtained from the intersecting point on the 4.0 AQL curve. Now reading TABLE X-K-1 below the chart, the corresponding 10% P_a level

is found in the left column. By reading across that row to the AQL column labeled "4.0", the value 12.3 is found. Note that this is a more precise estimate of the LQ obtained in the previous TABLE VI-A for a P_a - 10%.

There are many other definitions, plans, procedures contained in MIL-STD-105D. The instructor will aquaint the student/trainee with any additional aspects required for specific usage in his/her location.

PROCESS CONTROL

The basis of process control is the control chart which is used as an aid in determining if variation among the items is such as would occur in random sampling from a stable population. For instance, if successive samples of discharge water from a generating plant taken at hourly intervals over a long period of time resulted in a mean temperature of $22^{\circ}C$ with a standard deviation of 0.5°, the data could be used to construct a process control chart. Assuming the variation of water temperature followed a normal curve, additional samples taken at similar intervals should follow the same pattern. If a control chart such as Figure 27-7-11 is used in which the average of 22.1° is shown, with an upper control limit (UCL) and a lower control limit (LCL) at 3 standard deviations on either side of the mean, only about 3 plotted points per 1,000, on the average, would fall outside these limits.

$$(1 - .9973)\ (1000) = (0.0027)\ (1000) = 2.7 \approx 3 \text{ sample points}$$

In the above equation, remember that three standard deviations around the mean include 99.73% of the expected

samples measured (See Normal Curve section). Thus, when a point did fall outside the limits, there would be a check on the process immediately to determine if something has changed, if some explanation other than chance may be assigned. If the process is in control, "false alarms" would be obtained from 3 out of every 1,000 samples. Obviously, other control limits could be chosen to be more or less sensitive to possible changes in the process. Obviously process control does not tell whether the process itself is satisfactory, but only that it is under statistical control, and therefore, predictable.

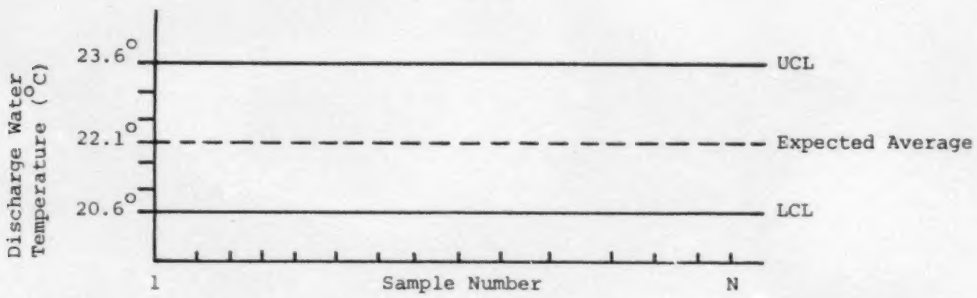


Figure 27-7-11 Upper and Lower Control Limits for Discharge Water Temperature

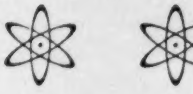
SUMMARY

This module presented the major points outlined below. Before proceeding to the Exercises, review each of the following major points.

1. The QA/QC Technician uses statistical techniques (a) to describe the characteristics of items such as size or weight; (b) to make decisions about acceptability of a batch of items from measurements made on a sample of items, and (c) to assess the accuracy and precision of the measuring instruments themselves.

2. No measurement is absolute due to variations from several factors including the measuring instrument and the individual making the instrument.
3. Measurements can be treated as absolute when assigned the mid-value of the measurement cell to which they belong.
4. An important characteristic of sets of data is the central tendency of the data as measured by the arithmetic mean ($\bar{x}$).
5. An important characteristic of a set of data is the dispersion of the measurements about the mean and this is measured by the standard deviation (s).
6. When large sets of measurements are made, the measurements often arrange themselves around the mean in a pattern that looks like a bell-shaped curve. The curve is also called the normal curve and it is useful in making decisions in both acceptance inspections and in process control operations.
7. Acceptance Inspection requires the selection of a sampling plan (MIL-STD-105D), which involves the taking of measurements on a lot sample and the statistical treatment of the measurements in order to accept or reject the lot.
8. Process control involves the taking of successive samples of the output of a system such as a pumping system, and applying statistical techniques to determine if the system is under control.
9. Any sampling plan involves two risks; one is the risk that a good lot will yield a bad sample (the producer's risk) and the second is that by pure chance a good sample may be drawn from a bad lot (the consumer's risk).
10. Sampling is based on randomness. The majority of sampling plans used in industry are obtained from the publication MIL-STD-105D.

11. There are four basic sampling plans encountered in practice which are based on an acceptable quality of 0.50 percent or fewer defective items in a lot. They are:
    a. Single Sampling
    b. Double Sampling
    c. Multiple Sampling
    d. Sequential Sampling



EXERCISES

In an effort to reinforce the concepts presented in this module, the following exercises have been subdivided into five sets with answer keys to each set.

Exercise Set No.	Title
1	True - False Questions on Overall Modules
2	Measurement Cell Exercises
3	Standard Deviation Exercises
4	Normal Curve Exercises
5	Statistics Exercises

1. True - False Exercise

(a) ____ Absolute measurements can be made by skilled technicians using accurate instruments.

(b) ____ A measurement cell represents an interval of uncertainty concerning true measurements.

(c) ____ Cell limits indicate the extreme upper and lower measurements possible within a measurement cell.

(d) ____ Cell intervals must be a constant length within one distribution.

(e) ____ The midpoint of a measurement cell is the most common measurement that falls within that cell.

(f) ____ Every measurement cell within a distribution will have a frequency of at least one.

(g) ____ Cell frequencies are obtained by tallying all the measurements within the distribution into the appropriate measurement cells.

(h) ____ A cell midpoint may be computed by adding the cell limits and dividing by 2.

(i) ____ A distribution always clusters around its arithmetic mean.

(j) ____ The mean is a common measure of central tendency.

(k) ____ The larger the standard deviation of a distribution, the more closely the measurements cluster around the mean.

(l) ____ A set of measurements that is normally distributed (or follows the normal curve) has many more measurements near its mean than at the extremes.

(m) ____ The total area under the normal curve is 100% of the distribution.

(n) ____ In a normally distributed set of measurements, 68.27% of the measurement will fall between two standard deviations below and two standard deviations above the mean.

(o) ____ Data that indicate whether the measurements are within or beyond some specified requirement are called attribute data.

(p) ____ The producer's risk is the probability that a bad lot will yield an acceptable sample.

(q) ____ The operating characteristic (OC) curve shows the risk involved with a particular sampling plan.

(r) ____ If cost were not considered, increasing the sample size would always be preferable.

(s) ____ When a product has a history of low acceptance, the Technician may be told to lower the acceptance standards as specified in MIL-STL-105D.

(t) ____ A control chart indicates whether a process is under statistical control.

Answers to True - False Exercises

(a) False
(b) True
(c) True
(d) True
(e) False
(f) False
(g) True
(h) True
(i) False
(j) True
(k) False
(l) True
(m) True
(n) False
(o) True
(p) False
(q) True
(r) False
(s) False
(t) True

2. Measurement Cell Exercise

(a) Determine the midpoints of the following measurement cells. Often you can determine the midpoint by inspection. You can also compute the midpoint of the interval (the arithmetic mean) by adding the cell limits together and dividing by 2.

(1)	10-20	(7)	225-275
(2)	500-1000	(8)	0.680-0.690
(3)	1.5-2.0	(9)	0.3325-0.3330
(4)	0.550-0.600	(10)	0.1020-0.1050
(5)	0.1050-0.1100	(11)	0.985-0.995
(6)	27.5-28.5	(12)	55.75-56.25

(b) A QA/QC person measured a sample of incoming parts and recorded the following measurements:

.4059	.4485	.4033	.4226
.4239	.4543	.4056	.4311
.4325	.4379	.4465	.4007
.4559	.4172	.4291	.4112
.4545	.4572	.4491	.4591
.4075	.4495	.4176	.4165
.4117	.4475	.4332	.4329
.4570	.4581	.4502	.4005

Complete the following table determining the cell midpoints and tallying the frequency for each measurement.

Hint: Frequency is the number of measurements that fall within a measurement cell.

The first measurement cell frequency and midpoint have been done for you.

Table for Listing Frequencies of Incoming Parts
in Specified Measurement Cells

MEASUREMENT CELLS	CELL MIDPOINTS	FREQUENCIES	TALLIES
.4550-.4600	.4575	5	~~𝍸~~
.4500-.4550			
.4450-.4500			
.4400-.4450			
.4350-.4400			
.4300-.4350			
.4250-.4300			
.4200-.4250			
.4150-.4200			
.4100-.4150			
.4050-.4100			
.4000-.4050			

(c) Using a cell interval of .25, set up a data table like the one in Exercise 2 for the following set of measurements:

.91	.56	2.21	2.23	1.73
1.22	2.41	.65	1.85	1.24
1.59	1.05	1.43	1.56	1.29
1.91	.88	1.62	2.19	1.39
.74	1.85	2.35	1.65	.71

Hints: Note the lowest and highest value above to establish the total range of numbers. Your lowest measurement cell should have cell limits of .50-.75. The midpoint values will have 3 decimal places (.625 for the measurement cell .50-.75).

Answers to Measurement Cell Exercises

(a)

(1)	15	(7)	250
(2)	750	(8)	0.685
(3)	1.75	(9)	0.33275
(4)	0.575	(10)	0.1035
(5)	0.1075	(11)	0.99
(6)	28	(12)	56

(b)

.4550-.4600	0.4575	5	~~1111~~
.4500-.4550	0.4525	3	111
.4450-.4500	0.4475	5	~~1111~~
.4400-.4450	0.4425	0	
.4350-.4400	0.4375	1	1
.4300-.4350	0.4325	4	1111
.4250-.4300	0.4275	1	1

.4300 .4250	0.4225	3	11
.4150-.4200	0.4175	3	111
.4100-.4150	0.4125	2	11
.4050-.4100	0.4075	3	111
.4000-.4050	0.4025	3	111

(c)

MEASUREMENT CELLS	CELL MIDPOINTS	FREQUENCIES	TALLIES
2.25-2.50	2.375	2	11
2.00-2.25	2.125	3	111
1.75-2.00	1.875	3	111
1.50-1.75	1.625	5	~~1111~~
1.25-1.50	1.375	3	111
1.00-1.25	1.125	3	111
0.75-1.00	0.875	2	11
0.50-.75	0.625	4	1111

3. Standard Deviation Exercises

(a) Use the information in the following data table to compute the standard deviation and mean of the measurements given.

Measurement Cell	Mid-Point (x)	Frequency (f)	f·x	x^2	$f \cdot x^2$
0.30-0.35	0.325	2	0.650	0.1056	0.2112
0.250-0.30	0.275	3	0.825	0.0756	0.2268
0.20-0.25	0.225	5	1.125	0.0506	0.2530
0.15-.20	0.175	5	0.875	0.0306	0.1530
0.10-.15	0.125	4	0.500	0.0156	0.0624
0.05-.10	0.075	1	0.075	0.0056	0.0056
Sums		N=20	$\Sigma fx=4.05$		$\Sigma fx^2=0.912$

Formulas:

$$s = \sqrt{\frac{N\Sigma fx^2 - (\Sigma fx)^2}{N(N-1)}}$$

$$\bar{x} = \frac{\Sigma fx}{N}$$

(b) Complete the data table below and determine the standard deviation of the measurements given.

Measurement Cell	Mid-Point (x)	Frequency (f)	f·x	x^2	$f \cdot x^2$
0.65-0.68	0.665	3	1.995	.4422	1.3266
0.62-0.65		5			
0.59-0.62		8			
0.56-0.59		6			
0.53-.056		2			
Sums		N=____	Σfx=____		Σfx^2= ____

$$s = \sqrt{\frac{N\Sigma fx^2 - (\Sigma fx)^2}{N(N-1)}}$$

(c) Set up a data table for the following set of measurements. Use cell intervals of 5.0. The lower limit of the first measurement cell should be 0. Use the formulas:

$$s = \sqrt{\frac{N\Sigma fx^2 - (\Sigma fx)^2}{N(N-1)}}$$

$$\bar{x} = \frac{\Sigma fx}{N}$$

to determine the standard deviation and mean of the measurements.

6.1, 7.2, 11.1, 19.8, 15.3, 18.0, 13.4, 9.7, 3.5, 18.3, 12.5, 7.7, 15.3, 10.5, 6.8, 14.2, 17.7, 4.9, 23.1

(d) A QA/QC Technician had instructions to reject a shipment of bearings if the standard deviation of the weights of the bearings was greater than 0.25 milligrams. Should the technician accept or reject the shipment of bearings from which the following weight measurements were taken?

9.5, 10.3, 9.9, 10.1, 9.8, 9.7, 10.4, 10.1, 9.5, 10.2, 9.8, 9.9, 10.2, 10.0, 9.9, 10.1 (All measurements in milligrams.)

Hint: Use 5 measurement cells starting with the lowest cell of 9.5-9.7.

Answers to Standard Deviation Exercises

(a) $s = 0.0695$

(b)

MEASUREMENT CELL	MID-VALUE (x)	FREQUENCY (f)	$f \cdot x$	x^2	$f \cdot x^2$
0.65-0.68	0.665	3	1.995	0.4422	1.3266
0.62-0.65	0.635	5	3.175	0.7669	3.9845
0.59-0.62	0.605	8	4.840	0.3660	2.9280
0.56-0.59	0.575	6	3.450	0.3306	1.9836
0.53-0.56	0.545	2	1.090	0.2970	0.5940
SUMS:		N=24	$\Sigma fx =$ 14.55		$\Sigma fx^2 =$ 10.8167

$$s = \sqrt{\frac{24\ (10.8167) - (14.55)^2}{24(23)}}$$

$$s = 0.2946$$

(c)

MEASUREMENT CELL	MID-POINT (x)	FREQUENCY (f)	$f \cdot x$	x^2	$f \cdot x^2$
20-25	22.5	1	22.5	506.25	506.25
15-20	17.5	6	105.0	306.25	1837.50
10-15	12.5	5	62.5	156.25	781.25
5-10	7.5	5	37.5	56.25	281.25
0- 5	2.5	2	5.0	6.25	12.50
SUMS:		N=19	Σfx= 232.5		Σfx^2= 3418.75

$$s = \sqrt{\frac{19\ (3418.75) - (232.5)^2}{19(19)}}$$

$s = 5.6455$

(d)

MEASUREMENT CELL	MID-POINT (x)	FREQUENCY (f)	$f \cdot x$	x^2	$f \cdot x^2$
10.3-10.5	10.4	2	20.8	108.16	216.32
10.1-10.3	10.2	5	51.0	104.04	520.20
9.9-10.1	10.0	4	40.0	100.00	400.00
9.7-9.9	9.8	3	29.4	96.04	288.12
9.5-9.7	9.6	2	19.2	92.16	184.32
SUM:		N=16	$\Sigma fx=$ 166.4		$\Sigma fx^2=$ 1608.96

$$s = \sqrt{\frac{16\ (1608.96) - (160.4)^2}{16(15)}}$$

$$\dot{s} = 0.2517$$

4. Normal Curve Exercises

(a) A supplier claims that the wire he produces has a mean breaking strength of 13.6 pounds with a standard deviation of 2.2 pounds. If this claim is valid:

(1) 95.45% of the wire sampled by a technician can be expected to have breaking strengths between _____ and _____.

(2) Only 3 out of each 1000 sample wires will have breaking strengths of less than _____ or more than _____.

(b) If a supplier of steel pins specifies that the pins have a mean weight of 12.4 ounces, with a standard deviation of 0.06 ounces, what percent of sample pins would a technician expect to weight a) between 12.34 ounces and 12.46 ounces? _____ (b) between 12.22 ounces and 12.58 ounces?_____

(c) The mean diameter of a lot of copper rods is claimed to be 16.3 millimeters with a standard deviation of 0.15 millimeters. Which of the following sample measurements are either too small or too large to fall within 99.73% of this distribution (±3 standard deviations)?

(1) 16.5
(2) 16.25
(3) 16.8
(4) 16.18
(5) 16.63
(6) 15.63
(7) 16.79
(8) 16.35
(9) 16.12

(d) Measure the major diameters of a box of 100 machine bolts (see Procedures section of this module) (UNC coarse or fine thread) using an outside micrometer and tabulate the data on the table in the Data Table section of this module. Using these data plot a Normal Curve for the distribution of bolt diameter and explain its shape.

Answers to Normal Curve Exercises

(a) (1) 9.2 and 18 pounds
(2) less than 7 or more than 20.2 pounds

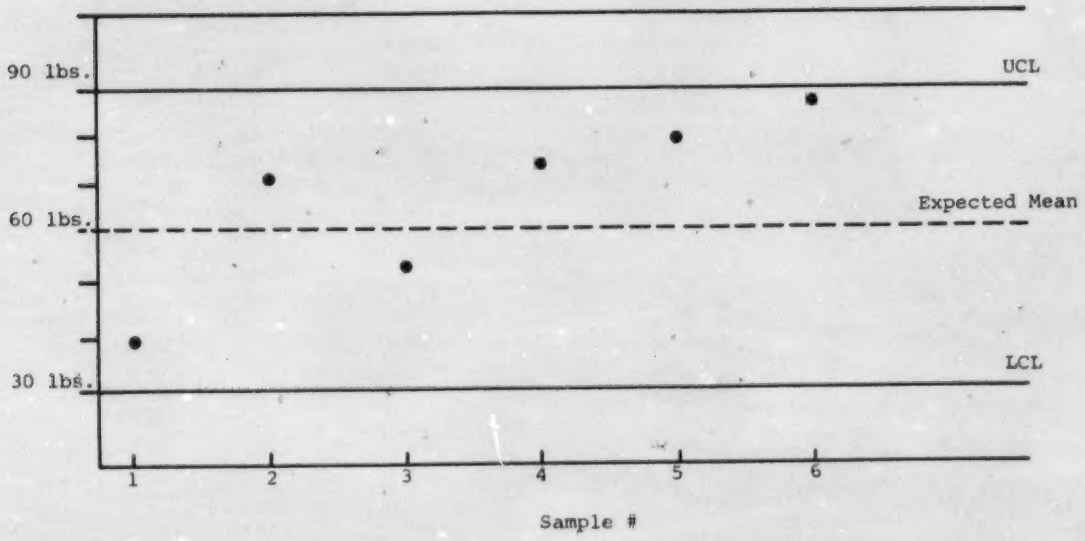
(b) (1) 68.27%
(2) 99.73%

(c) (1) 16.8 (too large)
(2) 15.63 (too small)
(3) 16.79 (too large)

(d) Refer to Procedures section of this module.

5. Statistics Exercise

(a) The following control chart records sample pressure readings of a storage tank.



(1) According to the six samples in the chart, is the process in control?

(2) What is the standard deviation that was used to set the UCL and LCL?

(3) How many pressure readings were above the mean?

(4) What reading was obtained on the 4th. sample?

(b) Construct a control chart for the following data obtained from repeated measurement of pollutants in samples of air. The expected mean is 50 particles per 1000 with a standard deviation of 5.

Sample #	Pollutants Per 1000 Particles
1	52
2	54
3	53
4	59
5	51
6	45
7	47
8	58
9	64
10	67

(1) Does the control chart assure the technician that the process is in control?

(2) Which sample reading(s) might lead the technician to check on the process?

Answers to Statistics Exercises

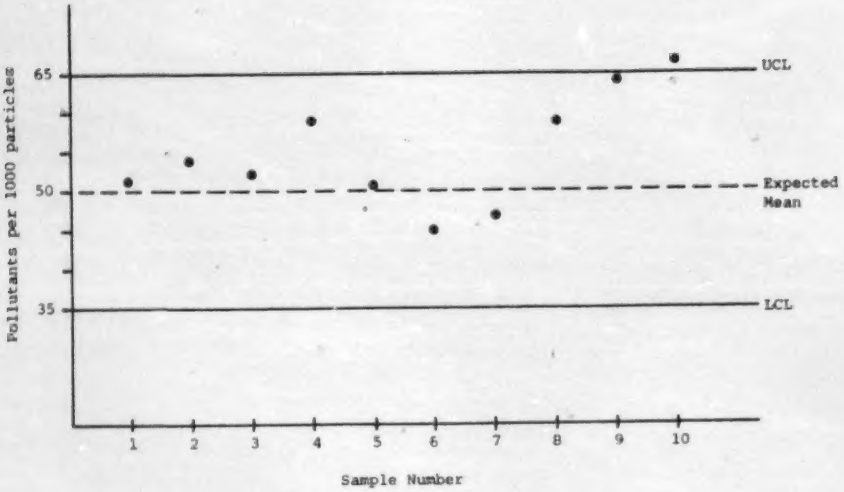
(a) (1) Yes

(2) 10 pounds

(3) 4

(4) 80 pounds (approximately)

(b) Plot:



(1) Yes

(2) Sample 10

MATERIALS

The following materials will provide the student and instructor with additional support in working on the Exercises and carrying out the Procedures listed in this module.

1. Hand Calculator
2. Quality Control textbook (see Reference Materials).
3. Outside Micrometer (one inch)
4. Box of machine bolts (minimum of 100 bolts of UNC) coarse or fine threads.

PROCEDURES

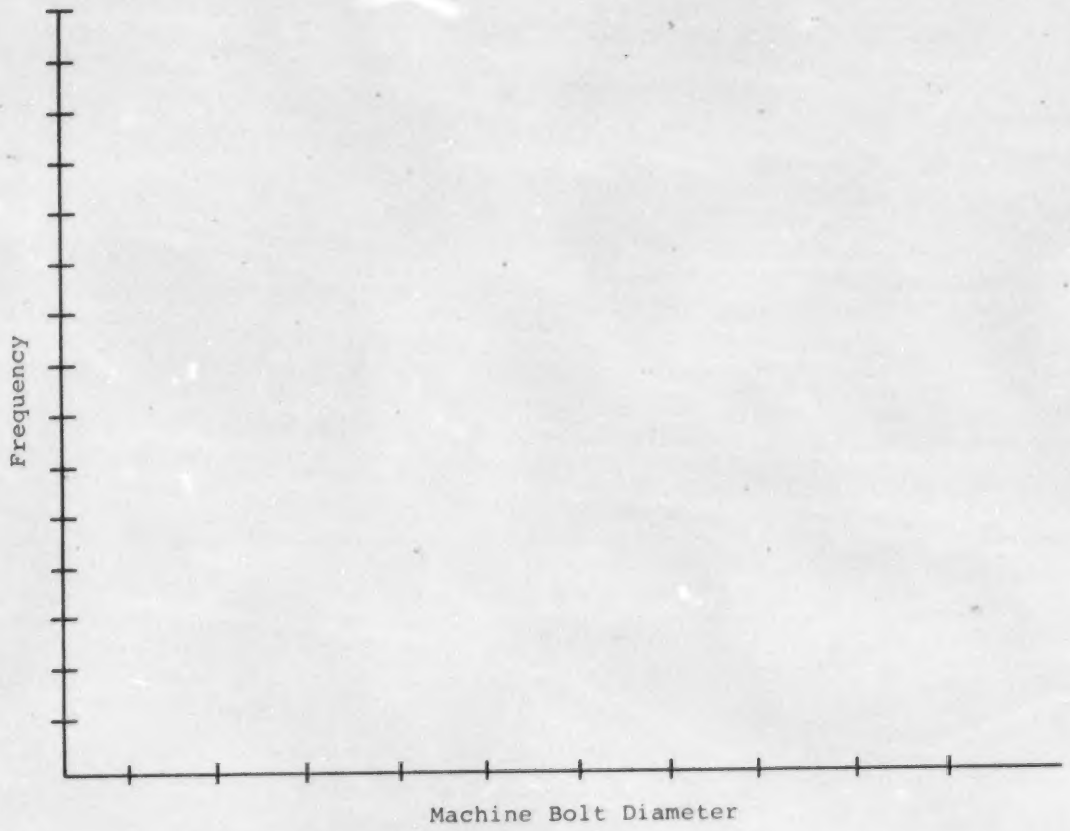
1. Measure the major diameter of a box of 100 machine bolts (UNC coarse or fine thread) using an outside micrometer.
2. Tabulate these data on the data table provided in the Data Table section which follows.
3. Plot the results on a graph of frequency (machine bolt number) vs. major diameter to obtain a normal distribution for these data. The graph is given in the Data Table section which follows.
4. Explain the shape of the normal distribution you have plotted and label the graph for one, two and three standard deviations.

Table 27-7-2
Machine Bolt Major Diameter

Machine Bolt Number	Major Diameter	Machine Bolt Number	Major Diameter	Machine Bolt Number	Major Diameter	Machine Bolt Number	Major Diameter
1		26		51		76	
2		27		52		77	
3		28		53		78	
4		29		54		79	
5		30		55		80	
6		31		56		81	
7		32		57		82	
8		33		58		83	
9		34		59		84	
10		35		60		85	
11		36		61		86	
12		37		62		87	
13		38		63		88	
14		39		64		89	
15		40		65		90	
16		41		66		91	
17		42		67		92	
18		43		68		93	
19		44		69		94	
20		45		70		95	
21		46		71		96	
22		47		72		97	
23		48		73		98	
24		49		74		99	
25		50		75		100	

Table 27-7-3

Graphical Plot of Data from Machine Bolt Major Diameter Measurement from Table 27-7-2



REFERENCE MATERIALS

Duncan, A.J. Quality Control and Industrial Statistics, Homewood, Ill.; Richard D. Irwin, Inc., 1965

Juran, J.M., Gryna, F.M., Jr., and Bingham, R.S., Jr. Eds., Quality Control Handbook. New York. McGraw Hill, 1974.

Military Standard. Sampling Procedures and Tables for Inspection by Attributes (MIL-STD-105D). Washington, DC.: Department of Defense, 29 April 1963.

TEST

Read each question carefully. Be sure to provide the specific information each question requires. Make your answers brief. Use the backs of the test if you need more space.

1. Determine the midpoint of the following measurement cells.
   a. 0.650-0.660
   b. 0.1050-0.1100
   c. 0.285-0.295
   d. 55.75-56.75

2. Determine the measurement cell limits for each of the following inspection activities:
   a. Measurement of 1.125 dimension to nearest 0.001 inch.

   _____
   _____
   _____
   _____

   b. Measurement of 0.1250 dimension to nearest 0.0005 inch.

   _____
   _____
   _____
   _____

   c. Measurement of 125.75 dimension to nearest 0.002 inch.

   _____
   _____
   _____
   _____

3. Using a measurement cell interval of 0.002 and 0.4460-0.4480 as lowest measurement cell, set up a data table showing measurement cells, frequencies, and cell midpoints for the following inspection results (use the table below for recording the data).

0.4478	0.4495	0.4555	0.4568	0.4514
0.4461	0.4501	0.4545	0.4508	0.4496
0.4483	0.4509	0.4573	0.4492	0.4525
0.4482	0.4525	0.4566	0.4502	0.4486

Measurement Cells	Frequencies	Cell Midpoints

4. Compute the mean of each set of numbers (give answers to one additional decimal place than used in each set)

a. 26, 32, 28, 42, 37, 33

b. 0.448, 0.450, 0.445, 0.458, 0.460, 0.442, 0.445

c. 0.025, 0.030, 0.018, 0.022, 0.024, 0.028, 0.029, 0.027

5. Complete the data table below and determine the standard deviation of the measurements given using the formula.

$$s = \sqrt{\frac{N\Sigma fx^2 - (\Sigma fx)^2}{N(N-1)}}$$

Give each answer to three decimal places.

Measurement Cell	Mid-Value x	Frequency f	f·x	x^2	$f \cdot x^2$
0.235-0.245		2			
0.225-0.235		3			
0.215-0.225		5			
0.205-0.215		5			
0.195-0.205		2			
Sums		N =	Σfx =		Σfx^2 =

Standard Deviation =

6. Describe each of the following sample plans.

a. Single sampling: _____

_____

_____

_____

b. Double sampling: _____

_____

_____

_____

c. Multiple sampling: _____

_____

_____

_____

7. Give one advantage and one disadvantage for each of the following sampling plans.

a. single sampling:

advantage: _____

_____

disadvantage: _____

_____

b. double sampling:

advantage: _____

_____

disadvantage: _____

_____

c. multiple sampling:

advantage: _____

_____

disadvantage: _____

_____

8. Describe producer's risk.

_____

_____

_____

_____

9. Describe buyer's risk.

_____

_____

_____